# *Synoptic preconditions for extreme flooding during the summer Asian Monsoon in the Mumbai area*

# Synoptic Preconditions for Extreme Flooding during the Summer Asian Monsoon in the Mumbai Area

MARCO LOMAZZI

*DHI Italia, Genoa, Italy*

DARA ENTEKHABI

*Parsons Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts*

JOAQUIM G. PINTO

*Department of Meteorology, University of Reading, Reading, United Kingdom, and Institute for Geophysics and Meteorology, University of Cologne, Cologne, Germany*

GIORGIO ROTH

*Department of Civil, Chemical and Environmental Engineering, University of Genoa, Genoa, Italy*

ROBERTO RUDARI

*CIMA Research Foundation, Savona, Italy*

(Manuscript received 28 February 2013, in final form 21 June 2013)

ABSTRACT

The summer monsoon season is an important hydrometeorological feature of the Indian subcontinent and it has significant socioeconomic impacts. This study is aimed at understanding the processes associated with the occurrence of catastrophic flood events. The study has two novel features that add to the existing body of knowledge about the South Asian monsoon: 1) it combines traditional hydrometeorological observations (rain gauge measurements) with unconventional data (media and state historical records of reported flooding) to produce value-added century-long time series of potential flood events and 2) it identifies the larger regional synoptic conditions leading to days with flood potential in the time series. The promise of mining unconventional data to extend hydrometeorological records is demonstrated in this study. The synoptic evolution of flooding events in the western-central coast of India and the densely populated Mumbai area are shown to correspond to active monsoon periods with embedded low pressure centers and have far-upstream influences from the western edge of the Indian Ocean basin. The coastal processes along the Arabian Peninsula where the currents interact with the continental shelf are found to be key features of extremes during the South Asian monsoon.

## 1. Introduction

The summer monsoon season is a distinguishing feature of the Indian subcontinent's hydrometeorology. Its arrival, evolution, and intensity have dramatic impacts on food security and exposure to hydrological hazards in the region. Monsoon rainfall provides about 80% of the annual precipitation over India. It dominates the region's water balance and is crucial for the agricultural sector. About two-thirds of the population is engaged in rain-fed agriculture and hence sensitive to this highly seasonal precipitation phenomenon. Rainfall events during the South Asia monsoon (SAM) season often have high intensities. In certain regions, daily rainfall totals above 200 mm are common. These intense rainfall events are likely to lead to flooding, with significant loss of life and economic damages. During the period 1980–2000, India had on average about 1300 human casualties

*Corresponding author address:* Giorgio Roth, University of Genoa, Via Montallegro 1, Genoa 16145, Italy.
E-mail: giorgio.roth@unige.it

DOI: 10.1175/JHM-D-13-039.1

per year associated with floods (Bureau for Crisis Prevention and Recovery 2004). This figure should partially be associated with the country's high population density. Growth and concentration of the population in denser urban settings is further exacerbating the problem.

Annual rainfall rates over India associated with the SAM are not spatially homogeneous. The zones with the highest precipitation are the western coastal strip and the easternmost part of the country, close to Bangladesh and Myanmar. Different studies recognize SAM complexity and highlight the need to focus on the study of monsoon regional characteristics through large-scale monsoon rainfall indices (Izumo et al. 2008; Vecchi and Harrison 2004).

The occurrence of SAM is associated with the intertropical convergence zone (ITCZ), which represents the ascending branch of the regional Hadley circulation (Charney 1969; Riehl 1979; Hartmann and Michelsen 1989; Gadgil and Joseph 2003; Gadgil 2003, and references therein). The Tibetan Plateau has recently been proven ineffective in driving large-scale SAM circulation, contrary to the Himalayas and adjacent mountain ranges (Boos and Kuang 2010). The ITCZ is located farther north during the Northern Hemisphere summer, affecting northern India between May and October, with maximum accumulated rainfall in July–August. In terms of upper-air conditions, the upper-air monsoon trough belt located over Tibet from June to September characterizes the SAM. The monsoon trough is associated with strong upper-air easterly flow (tropical easterly jet stream), whose upper divergence favors convection upstream of 70°E.

The SAM is characterized by rainy *active* periods and relatively dry *weak* periods, which give rise to intraseasonal variability (e.g., Oouchi et al. 2009) whose existence was first documented over a century ago (Blanford 1886). The intraseasonal variability is associated with the ITCZ fluctuations between two preferred latitudes: the northernmost one (associated with active periods) is located approximately between 15° and 25°N, while the southernmost one (associated with weak periods) lies close to the equator (Philander et al. 1996; Webster et al. 1998; Gadgil 2003; Ma et al. 2012). While the active and weak periods have been defined through synoptic conditions or through rainfall values, there is no unique definition in the literature (e.g., Rodwell 1997; Annamalai and Slingo 2001; Ma et al. 2012). Recent findings (Rajeevan et al. 2010) have shown how the active periods are associated with moist convection over the Indian Ocean (opposite pattern between the western and eastern parts of the ocean) whereas during prolonged weak periods a heat-trough-type circulation pattern toward India occurs (preceded by a meridional circulation over the Arabian Sea). For the Mumbai area, the monsoon onset is typically in the first half of June and monsoon withdrawal occurs in the second half of August (e.g., Wang and Lin 2002). Further, high rainfall days over western and southern India in June–August are linked to the moisture supply from the Somali low-level jet, while in September its relevance declines notably (Ordóñez et al. 2012).

As a first contribution, this study identifies extreme precipitation and flood events in the Mumbai region. From various sources and types of data, we establish a record of daily precipitation over the Mumbai region that spans one century, from which potential flood events are singled out. Such a long record is needed for the synoptic fingerprinting that is the second contribution of this study.

Given the fact that not all extreme precipitation events lead to flooding, and river discharge observations are sparse to nonexistent in the region, we develop an indirect method of inferring flooding events. The record of daily precipitation over the Mumbai region is combined with a unique unconventional dataset to infer a century-long estimated record of potential flooding. Over the Mumbai region, a catalog of flood events (recording at least the day, month, year, and location) has been assembled from state and local governments as well as media sources. This type of data (human reports rather than instrumentation measurements) poses challenges since the information may be biased. Nevertheless, their information content is important and without equal. In this study we use these reports to quantify the factors that make an extreme precipitation event a flood event.

Larger-scale and synoptic atmospheric conditions spanning the Indian subcontinent and neighboring oceans are contained in gridded atmospheric reanalyses datasets. We make conditional averages (also known as composites) of these synoptic fields by stratifying them into days with extreme precipitation and days with extreme precipitation and flooding; the difference between these two composites forms the fingerprints of hydrometeorological events that lead to flooding in the region.

## 2. Study area and data sources

As an economic center, the city of Mumbai is India's hub for finance, advertising, and media. With nearly 14 million inhabitants, it is also the second most populated city in the world (Hawksworth et al. 2009). The Mumbai region is located along the northernmost part of the western Indian coast, and approximately contained within a latitude–longitude rectangle bounded by 18°–20°N and 72.5°–73.5°E (Fig. 1). Relevant and synoptically effective surrounding regions include the Arabian Sea to the West and the Western Ghats mountain range to the east.

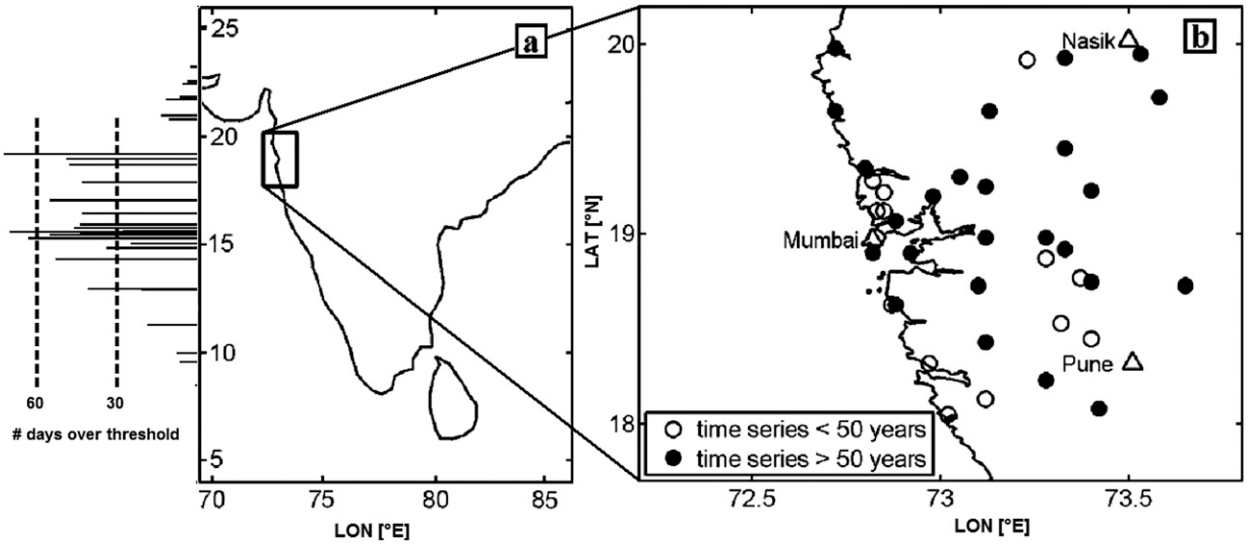


FIG. 1. (a) Study region (black rectangle) and number of days exceeding 15-mm daily precipitation accumulation during the 1951–87 period (histogram). (b) Location of precipitation gauges in the Mumbai area; the symbols distinguish the length of the available time series.

The histogram on the side of Fig. 1a shows that the Mumbai area is, along with Goa, often hit by extreme precipitation events associated with the SAM. This area experienced the intense event of 26 July 2005, which set a new historical record for daily precipitation accumulation over all of India (944 mm in 24 h recorded at the Santa Cruz airport rain gauge station). The related flood resulted in 1094 casualties in the city of Mumbai alone. Cattle losses exceeded 15 000 and the total economic loss reached almost $2 billion U.S. (Government of Maharashtra 2005).

Moreover, the daily maximum rainfall height over Mumbai for the 200-yr return period is expected to increase by a factor higher than 2 by the 2080s, excluding the 2005 event from the present-day estimation (Ranger et al. 2011). However, several coupled climate models used in the Intergovernmental Panel on Climate Change's (IPCC) Fourth Assessment Report (AR4) show a distorted representation of air–sea interactions in the Indian Ocean during boreal summer (Bollasina and Nigam 2009) and the two most recently upgraded ocean–atmosphere coupled model systems [European Centre for Medium-Range Weather Forecasts (ECMWF) system 4 and the National Centers for Environmental Prediction (NCEP) Climate Forecast System version 2 (CFSv2)] are deficient in predicting SAM precipitation (Kim et al. 2012).

### *a. Surface-based precipitation data*

To establish a quantitative definition of SAM extreme precipitation events, two rain gauge datasets are used: 1) long-term daily records of precipitation accumulation for a large number of stations during the period 1901–70 and 2) gridded precipitation data during the period 1951–2003.

Over India, daily data from approximately 3800 rain gauges spanning the period 1901–70 are available. Of these, 39 stations reporting daily precipitation fall within the study area. Out of the 39 stations, 25 have records for more than 50 yr. The remaining records have 15 yr or less of useful data (Fig. 1b). In the most interesting months (i.e., June–August), the mean number of available rain gauges is 27.

The extension of this database to include more recent years is made through the use of a gridded daily precipitation dataset. Gridded daily rainfall from 1951 to 2003 over all of the Indian territory is available through the Indian Meteorological Department (Rajeevan et al. 2006). The grid has $1° \times 1°$ cell size and covers the area included in the 6.5°–37.5°N and 66.5°–101.5°E box.

Since the 1951–70 period is covered by both datasets, a consistency analysis is carried out for those years. In Table 1 all the June–August daily accumulation statistical differences are divided into five ranges according to the rain amount. The root-mean-square error (RMSE) for each accumulation range is shown. RMSE values normalized by the median value in the range are included to characterize relative error in the accumulation ranges. The differences in the two datasets are small relative to the magnitude of the precipitation accumulation. With the two precipitation records now adjoined, the daily precipitation record for the Mumbai region spans more than a century, from 1901 to 2003. To extend further the daily precipitation records, it is possible to

TABLE 1. Comparison between rain gauges and gridded daily precipitation data: RMSE and error statistics. Different daily precipitation accumulation ranges are shown for the 1951–70 overlapping period.

| Station data vs gridded data | | |
|---|---|---|
| $P$ range (mm day$^{-1}$) | No. of days | RMSE (mm day$^{-1}$) |
| 0–5 | 370 | 1.3 |
| 5–50 | 1127 | 8.2 |
| 50–100 | 281 | 21.4 |
| 100–210 | 61 | 34.4 |

use satellite-derived precipitation, such as data from the Tropical Rainfall Measuring Mission (TRMM; Kummerow et al. 1998, 2000). However, the period of precipitation data is constrained by the availability of the unconventional flood reports and heritage reanalyses. Moreover, given the size of the region and the number of reporting rain gauges, the quality of the surface data is adequate and possibly even superior to satellite-based records in capturing small-scale features. The gauge-recorded daily precipitation accumulation also avoids the issue of diurnal sampling biases in tropical regions. Bias correction is usually needed for satellite-based products (Bowman et al. 2005; Sanderson et al. 2006).

### b. *Unconventional data: Flood reports*

A unique aspect of this study is the use of unconventional data to create added-value products. A comprehensive historical record of reported flood events (human reports by agents in the local or federal government or by the media) across India has been made available for this study (Munich Re Group 2008a,b). This flood catalog reports on approximately 830 flood events over the entire India territory during the period 1900–78 in the months from May to December. Information about the events includes the date (sometimes only the year and month), the path followed by the perturbation, the kind of weather system (when available, distinguishing among tropical storms, tropical depressions, and tropical cyclones), and the damage caused (when available and including casualties, economic losses, and livestock losses).

Over the Mumbai region, 27 flood events are identified in the flood catalog during the June–August period. These examples are marked as confirmed flood events (CFEs). Because these are human reports, they are prone to biases. In fact, the reporting of flood events is, among other factors, mediated by the presence of observers, the availability of information exchange media and opportunities, and incentives to underreport or overreport. Floodwaters that do not cause damage may go unreported,

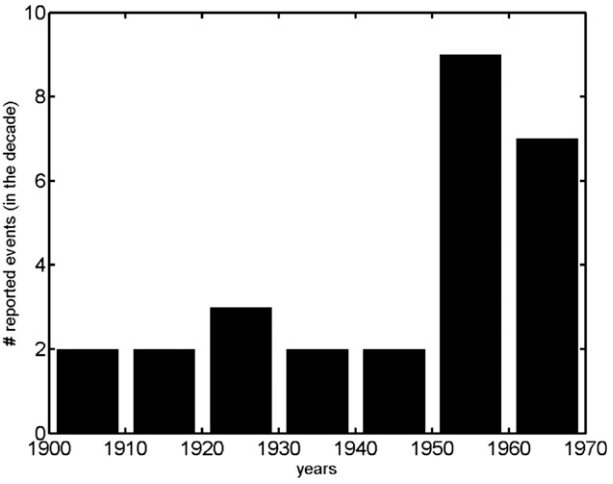


FIG. 2. Number of CFEs in the Mumbai area reported in the flood catalog.

while less extreme flood events over more vulnerable settlements may be overreported. Furthermore, the increase in both population and urbanization leads to increased vulnerability.

These aspects are evident in Fig. 2, which is a count of CFEs for each decade in the flood catalog. The last two decades have by far a higher number of CFEs (averaging eight cases per decade) than the previous decades (around two cases). The above-mentioned factors are key in determining such discrepancies. Inferring trends from the raw flood catalog information can, therefore, be misleading.

Working with unconventional data has the hazards listed above and more. Nonetheless, if the information content can be extracted and used to add value to instrumental records, then unique opportunities exist for the long-record analysis of hydrometeorological phenomena and natural hazards. In this study we use the unconventional data as a guide to delineate thresholds in instrumental measurements to define potential flood event (PFE) days. This is performed by using the CFEs to delineate—in a precipitation accumulation versus soil moisture space—the zones that define potential flood events, that is, to reconstruct a century-long record of days with likely flood events. The approach is described in detail in section 3a. Finally, the Emergency Events Database (EM-DAT; www.emdat.be) is used to check the consistency of the CFEs database across the reporting period. From 1900 to 1970, the Em-DAT database reports 18 events that could be related with the study area against the 21 reported as CFEs. Unfortunately, during this period EM-DAT rarely reports the geographic location of the event and in many cases the date of the event is approximate. With such information,

only a qualitative consistency check may be carried out, which does not indicate the presence of relevant misses in the CFE catalog.

### c. Atmospheric data for synoptic analysis

Two different datasets are combined to characterize the synoptic conditions that lead to extreme precipitation and flood events. Surface precipitation gauge records are used as indices to sample gridded atmospheric reanalysis data and eventually to generate composite synoptic maps. Within this framework, reanalysis data cover a much larger region and provide the synoptic setting leading to the events.

In this study we use the 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40; Uppala et al. 2005). The data are available from 1958 until 2002 (inclusive) at 1.125° resolution. We specifically use the mean sea level pressure (MSLP), horizontal wind vector, and humidity flux vector (defined as the product between horizontal wind components and specific humidity).

## 3. Definition of flood events

Heavy precipitation and antecedent surface wetness are cofactors in producing flood events. Neither factor alone necessarily leads to floods. Starting from this rationale, the CFEs and the extensive daily precipitation records are used to reconstruct a proxy record of flood events over the Mumbai region that spans the twentieth century.

### a. Antecedent surface wetness and heavy precipitation

The reconstruction of daily flood events begins with developing estimates of antecedent surface wetness. To this aim, station records are weighted by area (Thiessen polygons or the nearest-neighbor method) to obtain a representative time series of the average daily precipitation over the study area. The June–August daily rainfall data recorded at the 39 rain gauges shown in Fig. 1 for the period from 1901 to 1970 are used for this purpose.

The average daily precipitation totals are then filtered using the antecedent precipitation index (API), a one-sided filter precursor to more complex models such as soil moisture accounting and soil moisture diffusion models. API is used here as a parsimonious and adequate index of the surface wetness cofactor in producing floods from precipitation events. Within this context, more parameter-heavy land surface models are not suitable.

The API index accumulates daily precipitation totals and dissipates the accumulation by a dissipative term (fraction of the index) $k$ as in

$$\mathrm{API}_d = k \times \mathrm{API}_{d-1} + P_{d-1}, \quad (1)$$

where $\mathrm{API}_d$ is the API of day $d$; $\mathrm{API}_{d-1}$ and $P_{d-1}$, respectively, are the API and the precipitation total of the day before; and $k$ is the dimensionless decay coefficient. The magnitude of the linear decay term depends on a number of factors including depth of the effective hydrologic soil layer, time step of the model, soil texture, rooting and vegetation cover, and seasonal climate. There is an extensive history of experiments and applications that address the magnitude of the parameter in different settings. Studies can be traced back to Saxton and Lenz (1967) and Blanchard et al. (1981) among others. The value of the linear decay generally lays in the narrow 0.92–0.99 range with 0.93 the most common nominal value. In this study we also adopt the 0.93 value based on past experiences with the implementation of the model.

The two flood cofactors (API and $P$) for each day of the 1901–70 record are shown in Fig. 3a. On the basis of the flood catalog dataset, a specific day can be associated with every CFE event as the day with the highest rainfall total within the duration of the flood event. Figure 3a shows that indeed the CFEs are associated with high values for both cofactors: wet antecedent surface conditions and heavy precipitation totals. Based on this analysis, CFEs can now guide the definition of the API and $P$ thresholds used to develop a long proxy record of flood events that goes beyond the flood catalog reports.

### b. $P$ and API thresholds

Based on their marginal probability distributions, $P$ and API are first transformed into dimensionless percentiles, as shown in Fig. 3b. Threshold values for PFE identification are then selected to envelope the CFEs as much as possible, with reasonable allowance for outliers in the flood reports. The result is the threshold value 0.98 for the precipitation total percentile (78 mm day$^{-1}$) and 0.78 for API (209 mm). Accordingly, a PFE is defined as a day in the record when both the P and API values are equal or higher than the twin thresholds. In total, 5 CFEs are not considered to be PFEs, and 165 (13% of the original 1280) rainfall events are included in the PFE dataset.

### c. PFEs in the extended record

Data used to define the $P$ and API thresholds are based on the stations' daily precipitation. These thresholds are applied to the extended record, augmented with gridded data as described in section 2a. As a test of consistency in PFE detection between the two precipitation datasets (station and gridded), the overlapping

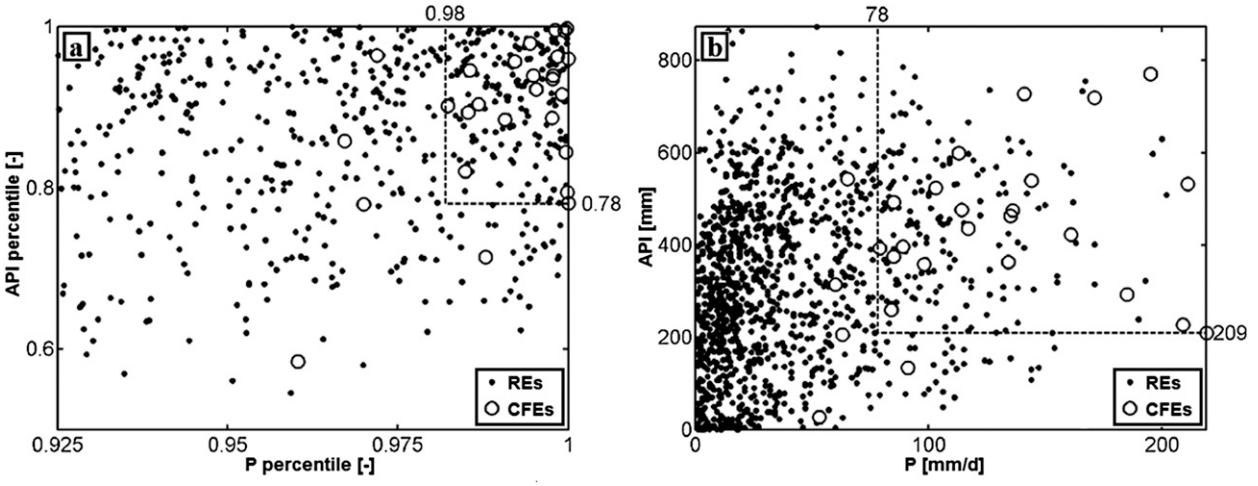


FIG. 3. Pairs of daily API and daily precipitation totals for the study region during the 1901–73 period (black dots). CFE days are marked by empty circles. CFEs confirm that flood event days are associated with high values in both cofactors for floods: wet antecedent surface conditions and heavy precipitation. Dashed lines define the region in the API–*P* space associated with floods and contain most of the CFEs. The data are shown (a) as percentiles in their respective marginal distributions and (b) as API and *P* values in units of accumulation and daily rate (mm and mm day$^{-1}$).

period is investigated. First, the percentile thresholds for *P* and API are translated to accumulation values in each dataset. Then, PFEs are identified using the twin-threshold method for the overlap period and the number of PFEs identified in both datasets and those identified in one, but not the other, counted. Table 2 shows the results of this consistency check. Most PFEs are identified in both datasets: less than 20% are detected in one but not the other. This adds confidence in the reconstruction of the century-long PFEs time series obtained using the augmented daily precipitation records.

Figure 4 shows the plot of the precipitation accumulation and antecedent surface wetness index for the entire century-long record, from 1901 to 2003. Applied PFE thresholds are depicted by the dark-gray region in the top-right corner. Figure 5 shows the resulting number of PFEs in each decade of the twentieth century. The number of cases is stationary and varies between 20 or 30 per decade, that is, approximately 2–3 per year. Whereas the CFEs show the biases expected from these kinds of reports (see Fig. 2 and the related discussion), the reconstructed PFEs are more objective and they do not show significant trends.

EM-DAT and Dartmouth Flood Observatory (DFO, http://floodobservatory.colorado.edu) databases have been used to check PFE identification for 1985–2003 (DFO is only available from 1985). During this period, the EM-DAT database reports dates more reliably, so that a better comparison is possible. EM-DAT reports 24 events whose dates are included in the 45 identified PFEs. Among these, 6 can be precisely located in the study area. DFO reports 11 events over the study area. All of them are included in the set of PFEs.

### *d. High rainfall and soil moisture events (HRSEs) as the benchmark set of events*

To detect the synoptic fingerprint of flood events in the long record, a ''baseline'' case needs to be defined. The baseline case is the composite of regional atmospheric fields for precipitating monsoon days not identified as PFEs. Based on this rational, we define another set of thresholds for both *P* and API to identify the benchmark set of events: the high rainfall and soil moisture events (HRSEs). Threshold values should serve the function of the study: the closer the values to the PFEs thresholds, the more comparable are the baseline case

TABLE 2. Number of PFEs in the 1951–70 overlapping period selected by both datasets or by one of them.

| | | 1951–55 | | 1956–60 | | 1961–65 | | 1966–70 | |
|---|---|---|---|---|---|---|---|---|---|
| PFEs database | *P* threshold (mm day$^{-1}$) | No. of PFEs | Mean *P* (mm day$^{-1}$) | No. of PFEs | Mean *P* (mm day$^{-1}$) | No. of PFEs | Mean *P* (mm day$^{-1}$) | No. of PFEs | Mean *P* (mm day$^{-1}$) |
| Grids | 62 | 2 | 63 | 3 | 67 | 3 | 69 | 2 | 71 |
| Stations | 78 | 2 | 83 | 1 | 94 | 5 | 88 | 1 | 79 |
| Both | | 13 | | 17 | | 15 | | 13 | |

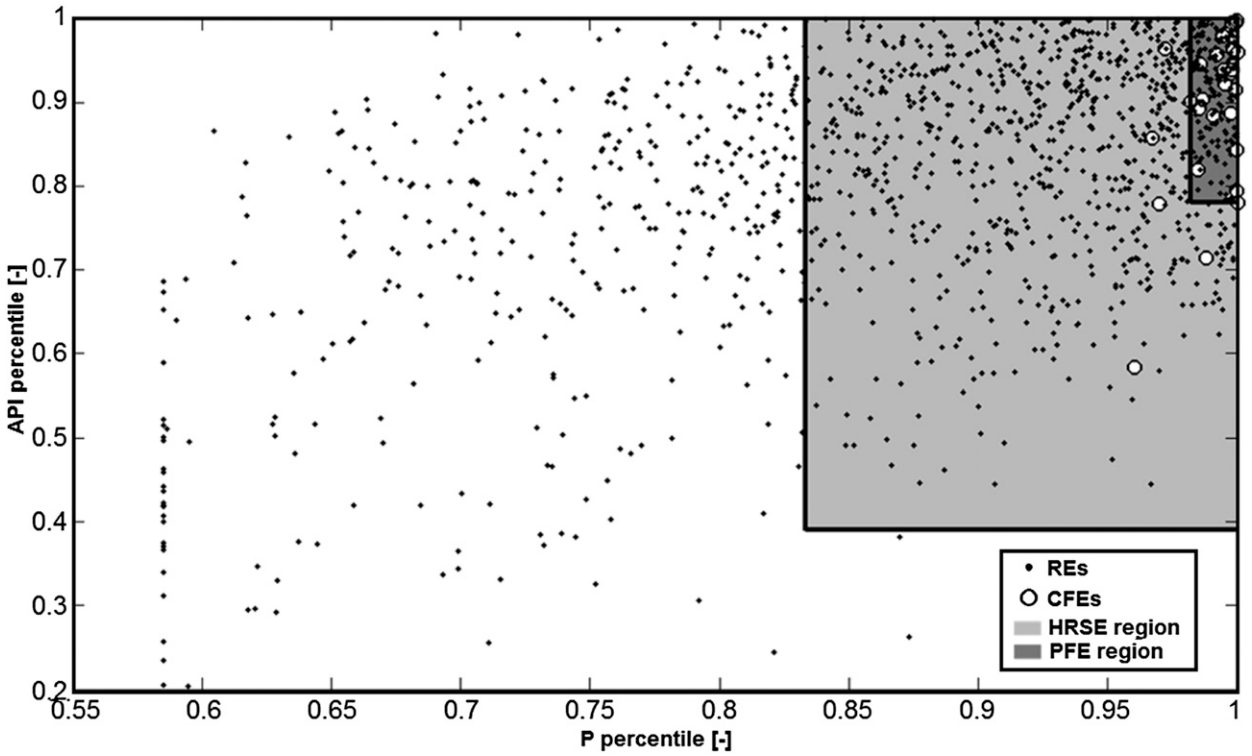


FIG. 4. PFEs and HRSEs regions in the 1901–2003 extended record. CFEs and REs are shown as empty circles and black dots, respectively. Axes are in *P* and API percentile values.

and the PFEs. On the other hand, lower thresholds would both identify a larger number of samples for the statistical pool and increase the possibility of including more local precipitation events (e.g., those associated with the Western Ghats). We selected the 0.833 and 0.390 percentiles for *P* and API, respectively. A total of about 900 HRSEs exist in the record. The days with PFEs and HRSEs are next used to develop regional atmospheric field composites.

## 4. Dynamic and synoptic analyses

### *a. Composites of atmospheric fields associated with* PFE*s and* HRSE*s*

We start our synoptic diagnoses by analyzing MSLP fields corresponding to the period of the rainfall events in the database. Composites of regional MSLP are produced for PFE and HRSE days, respectively. The difference between these two composites is indicative of the key characteristics of SAM episodes that lead to flooding over the region. Since the ERA-40 data span the years 1958–2002, only events in this period have been considered for detailed analysis: 125 PFEs and 454 HRSEs are identified.

Both PFEs and HRSEs are associated with a low pressure area centered over northwestern India and southern Pakistan, but PFE minima (around 995 hPa) are approximately 3 hPa deeper than HRSE cases. This synoptic situation corresponds to active monsoon conditions with low pressure systems, which play an important role in the monsoon rainfall (Krishnamurthy and Ajayamohan 2010). Over the Mumbai area, the HRSE

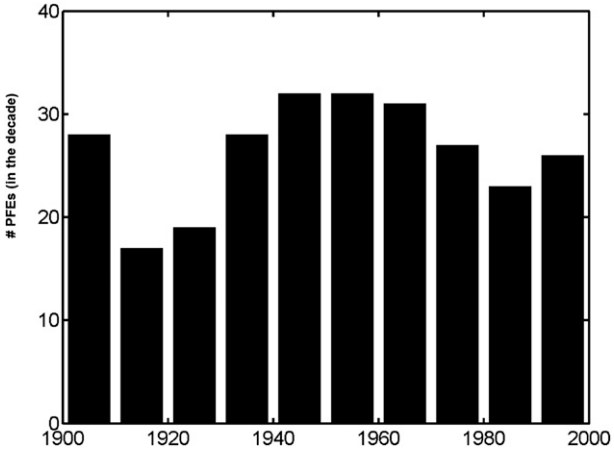


FIG. 5. Number of PFEs in each decade from 1900 to 2000.

pressure values remain practically constant (slightly larger than 1002 hPa), while PFE values decrease from around 1001 hPa 2 days before the event day to 999.5 hPa on the event day. In both cases, the low pressure areas are associated with the upper-air monsoon trough. The presence of low pressure systems and the upper-air trough leads to a large-scale ascent of air over the region, supporting the triggering of convection and the formation of precipitation (cf. Krishnamurthy and Ajayamohan 2010). Figure 6 shows the difference between the MSLP fields associated with PFEs and HRSEs, from 2 days before to the event day. The dotted regions have statistical significance equal or higher than 99% in the mean of the two samples. The difference field shows the pattern of the PFE evolution over the Mumbai region. Two days before the potential flood event the maximum difference is located over the northern part of the Bay of Bengal, off the coast of Orissa. In that region, the MSLP values associated with PFEs is on average between 997 and 998 hPa, approximately 2.5 hPa lower than HRSE values. Two days before a PFE, the MSLP over the study area is approximately 1.5 hPa lower than for HRSE monsoon days. One day before the event, MSLP over the Mumbai area is about 2 hPa lower for PFE events compared to HRSE events. The maximum difference is still located over the coastal part of Orissa and is slightly higher than 2.5 hPa. On the day of the event, the highest MSLP difference (nearly −3 hPa) is located a few degrees east of the study area, over which the PFE value (approximately 999 hPa) is around 2.5 hPa lower than that for the HRSEs. The MSLP signature confirms the importance of the larger field intensity and more importantly the location of cyclonic pressure systems in inducing flood conditions over the region. The pressure patterns and their evolution drive atmospheric fluxes and flux convergences that lead to flood-producing persistent and intense precipitation events. The trajectory of the moisture fluxes is key to such flooding episodes.

### b. *Back-trajectory analysis of air parcels*

To better understand the regional synoptic preconditions for precipitation episodes that lead to floods, the trajectories of the air parcels carrying moisture toward the study area are analyzed. Back-trajectory analyses for flood cases are usually inconclusive since each event has its own unique conditions. Nevertheless, the large statistical sample of events in this study (125 PFEs and 454 HRSEs) allows us to develop mean trajectories with tests of statistical significance. Figure 7 shows the density of trajectories from −72 h until the event day (contours) and the location of particles 72 h before the event. The pressure levels considered are 850 and 700 hPa. These

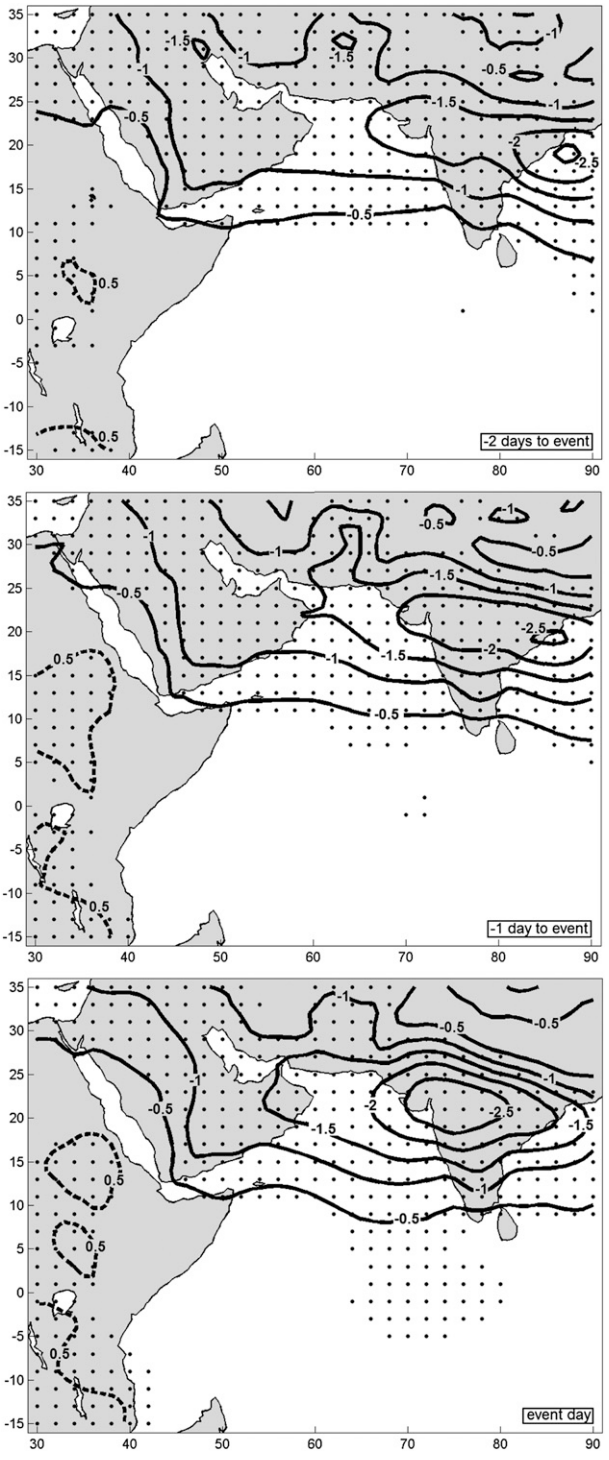


FIG. 6. Difference between PFE and HRSE MSLP composites (contours) from (top) −2 days to (bottom) the event day. Black dots show regions with ≥99% statistical significance in the difference metric. Dotted contours are positive differences and solid contours are negative differences.

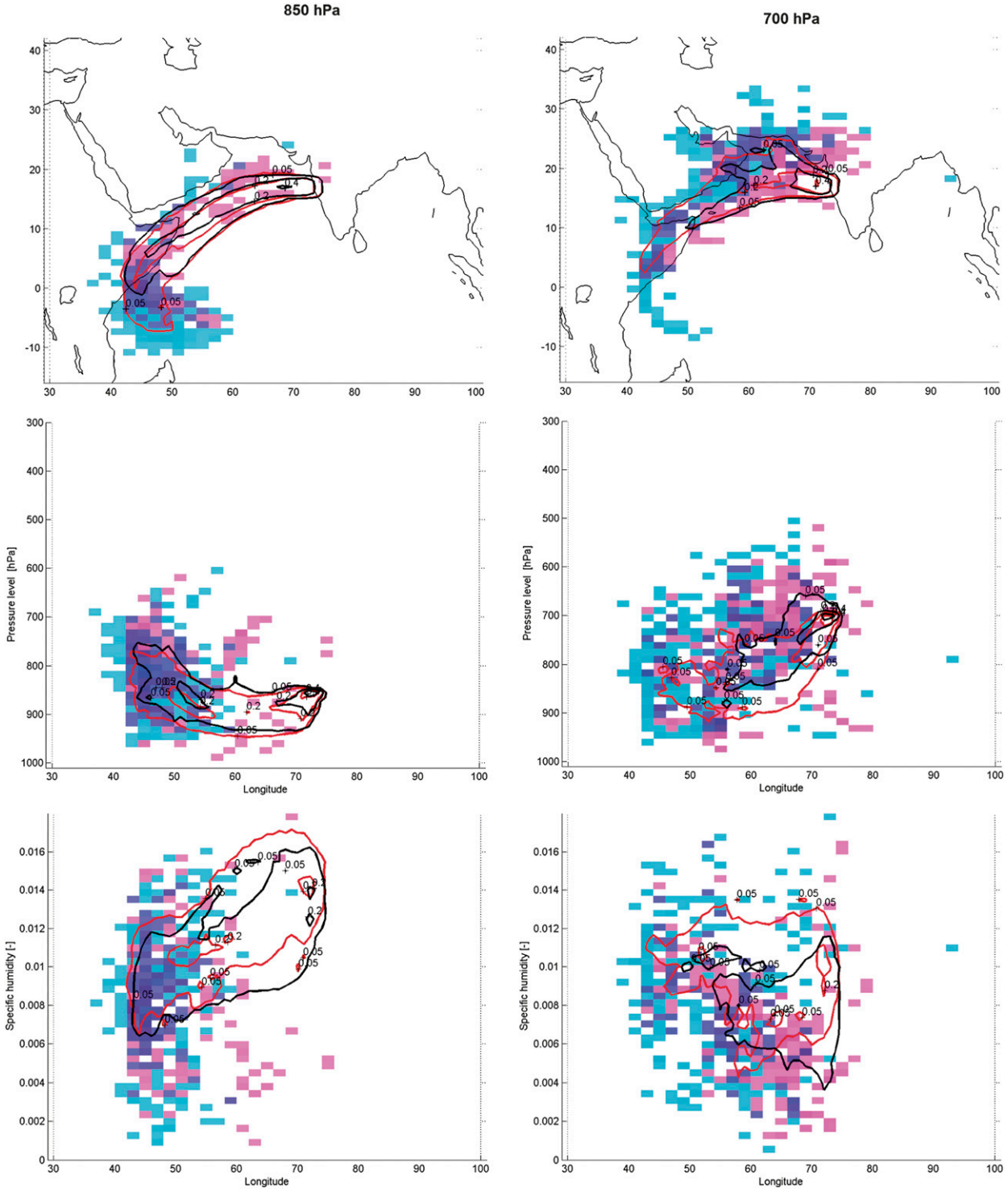


FIG. 7. Back-trajectory density contours of PFEs (red contours) and HRSEs (black contours). Cyan shading represents the starting points of the PFE back trajectories, magenta shading represents the starting points of the HRSE back trajectories, and light blue areas are common starting points of back trajectories: (left) 850 and (right) 700 hPa. (top) latitude-longitude planimetric view, (middle) elevation-longitude view, and (bottom) longitudinal profile of specific humidity.

are the levels at which most of the moisture flux is concentrated. In addition, the fields (planimetric and vertical) of the specific humidity along the longitude of the trajectory are analyzed in order to track the evolution of the flux along the trajectory (Methven 1997; Methven et al. 2001).

The distinctions in the trajectories associated with PFEs and HRSEs are shown in Fig. 7. The trajectories at 700 hPa are more diverse than those for 850 hPa. This characteristic holds for the planimetric and vertical locations of moist particles and for the specific humidity panels as well. The main reason is the higher variability of the wind trajectories at medium levels than at lower levels for the two case samples. At 72 h before the event day, the 700-hPa particles (right panels in Fig. 7) are more scattered than those at 850 hPa (left panels). This is especially true for PFEs particles. As the top panels in Fig. 7 show, 72 h before the event nearly all 850-hPa PFE particles are clustered between Madagascar and Somalia, while 700-hPa PFE particles lie over a bowlike area from Somalia to the easternmost part of India. The 850-hPa particles follow a clear path from off the southern coast of Somalia 3 days before the event. Such a path is characteristic of the Somali low-level jet, which previous studies show to be a primary source of moisture for western and southern India during SAM (Zhou and Yu 2005; Jinhai et al. 2007; Ordóñez et al. 2012). This corresponds to the typical conditions of the active monsoon, when enhanced moisture transport is found toward the study region.

Particles move northward around the Horn of Africa and then veer north-eastward, passing over the Arabian Sea to reach the Mumbai region. When over eastern Africa, the vast majority of particles are between 900 and 700 hPa. Over the central Arabian Sea their height decreases to around 900 hPa while in the very last part of the path they rise to 850 hPa. While passing over the Arabian Sea, the particles considerably increase their moisture content by picking up moisture over the ocean (PFE particles nearly double their specific humidity).

The PFE and HRSE 850-hPa particles are most distinct when their planimetric location and humidity content are considered. HRSE particles on average precede PFE ones: 72 h before the event a fraction of HRSE particles are located over the Arabian Sea (some of them are even close to the Mumbai coast) and in the ensuing days most HRSE particles maintain positions ahead of their PFE counterparts. This is mainly due to distinctions in wind speed, which is lower over the Arabian Sea in the case of HRSEs (Fig. 7).

The two particle samples are differentiated with regard to their moisture content while passing over the Arabian Sea. At 3 days before the event, HRSE and PFE particles do not show a clear distinction in their specific humidity values (mostly in the 0.007–0.012 range). Once they reach the study area most PFE particles are in the range of 0.011–0.017, while HRSE examples are in the range 0.009–0.015. The difference peaks over the central Arabian Sea at 65°E. At this location the PFE range is 0.01–0.017 and the HRSE range is 0.008–0.014. The HRSE particles tend to rapidly increase their moisture content a few degrees longitude west of the study area, while PFE particles accumulate the greater part of their humidity before reaching the central Arabian Sea.

Particles at 700 hPa show a different pattern of behavior compared with those at 850 hPa. In the horizontal plan a clear path for 850-hPa particles is not clearly discernible. But the trajectories of both PFE and HRSE particles converge from starting points scattered from the Horn of Africa to the India subcontinent, leading to the study area. On the day before the event, the 700-hPa particles are closer to the Mumbai area are than 850-hPa ones. As particles are aligned in the southwest–northeast direction, PFE particles tend to be located north of HRSE ones along this direction.

The specific humidity of 700-hPa particles has a lower value than that of the 850-hPa examples. In comparison with 850-hPa particles, the difference between the PFE and HRSE particle moisture content is more distinct. A few degrees east of the Mumbai area, most 700-hPa PFE particles have a specific humidity in the range 0.006–0.013 while most HRSE cases are in the range 0.004–0.01.

The trajectories of moisture leading to floods in the Mumbai region originate along the eastern coast of Africa and the Arabian Peninsula. Remarkably, the originations and paths are confined in narrow corridors. This is consistence with the presence of an intense low-level Somali jet. The waters off the southern Somali coast through the Arabian Sea are key to the flood episodes over the region.

### *c. Wind patterns and moisture flux*

The horizontal wind (direction and speed) and the moisture flux fields are backdrops for the particle trajectories. The composites of horizontal winds at 850 hPa for PFEs and HRSEs are estimated for the larger area surrounding the study region. The focus is on the lower-atmospheric levels where an atmospheric moisture pathway is present. Figure 8a shows the mean PFE wind direction (arrows) and the difference between PFE and HRSE composites of the wind component along the mean PFE direction (contours) on the event day. Arrows are shown only where the statistical significance of the PFE and HRSE difference is at least 99%.

The patterns confirm that the mean PFE wind direction coincides with the moist particle paths identified

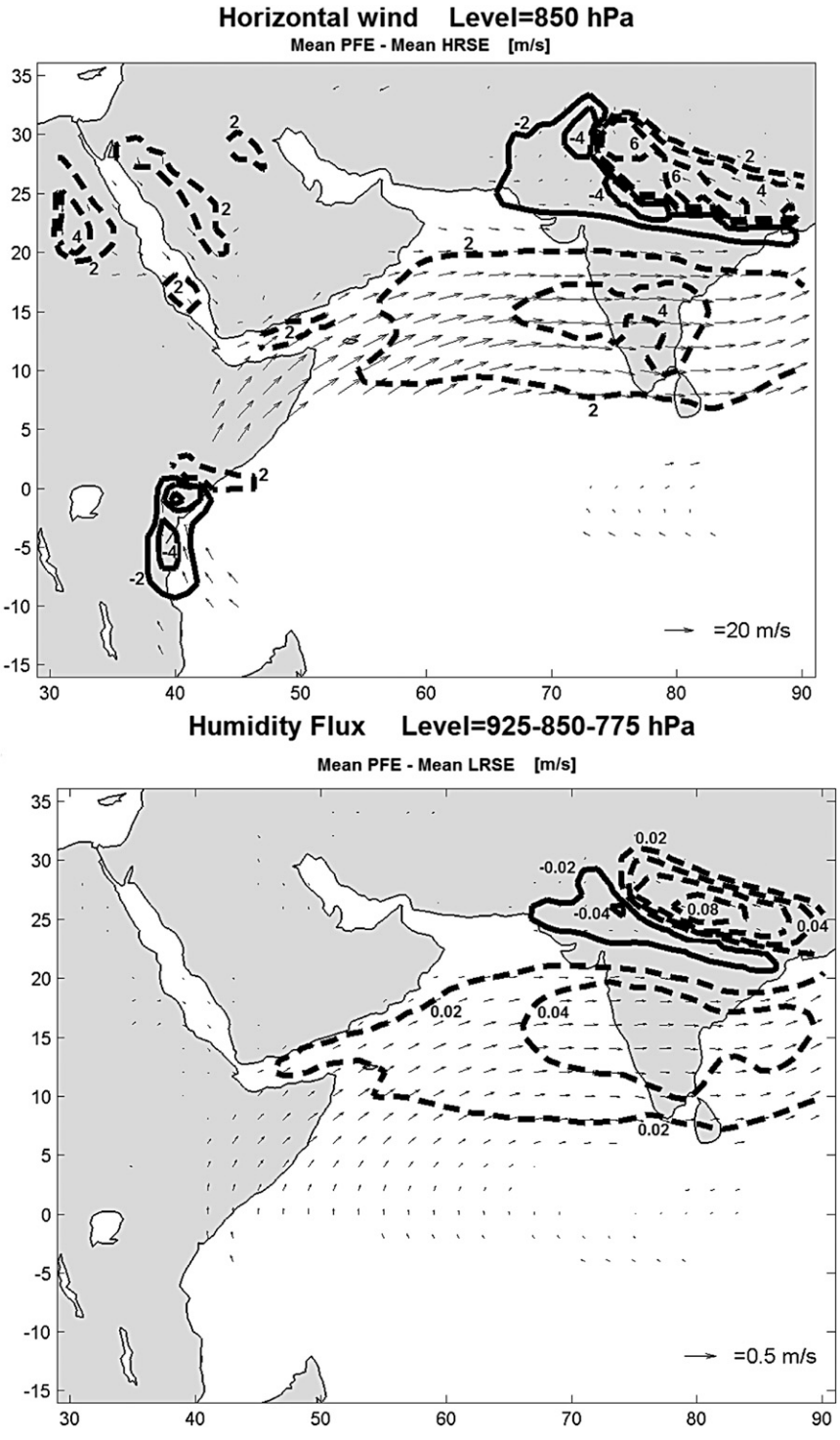


FIG. 8. (a) Mean horizontal wind anomaly patterns (i.e., difference in PFE and HRSE composites) along the mean PFE wind direction for the event day. (b) Humidity flux anomaly patterns along the mean PFE wind direction for the event day. Vectors show the mean PFE wind direction and speed and are drawn only if they are above the 99% statistical significance anomalies.

with the trajectory analysis. They originate at the western boundary of the Indian Ocean at approximately 15°S. The airstream moves northwestward to the oceanic area between Somalia and Madagascar, then it veers northeastward over Somalia and western Arabian Sea and finally over the central Arabian Sea. This confirms that the Somali low-level jet is reinforced, in agreement with conditions associated with an active monsoon period. The mass flux then turns into zonal flow on its path to the Mumbai area (850-hPa wind velocity of around $20\,m\,s^{-1}$). As also shown in the trajectory analysis, the PFE horizontal wind speed is higher than in the HRSE cases. Over the Arabian Sea and the Bay of Bengal the difference exceeds $2\,m\,s^{-1}$, with peaks exceeding $4\,m\,s^{-1}$ over peninsular India.

Another notable difference between the PFE and HRSE atmospheric patterns is in continental India. Over its central and eastern parts, the PFE wind speed exceeds HRSE winds by up to $6\,m\,s^{-1}$. In contrast, the western and southern parts of continental India are characterized by HRSE wind speeds that are higher than those in PFE composites. The differences can be as large as $4\,m\,s^{-1}$. The airstream hits continental India after crossing peninsular India and subsequently veers over the northern Bay of Bengal. The mass flux then traverses continental India and moves northward through the Punjab and Kashmir regions. Negative values over western continental India are due to the location farther south of the center of the PFE cyclonic circulation (23°N, 80°E). The pattern observed in Fig. 8 is similar to active monsoon conditions with embedded low pressure centers, as described by Krishnamurthy and Ajayamohan (2010, their Fig. 8a). The weaker HRSE composite cyclonic circulation is centered at 26°N, 81°E. The PFE circulation is characterized by high wind speeds north of the cyclone center and, importantly, confined in a channel bounded to the north by the Himalayan mountain range. The HRSE composite circulation is characterized by higher wind speeds west of the cyclone core.

The increased circulation over the Arabian Sea, unlike the winds pattern over continental India, is a large-scale atmospheric pattern that begins at least from 3 days before the event day (not shown). It plays an important role in triggering PFEs over the study area. The circulation at lower-atmospheric levels picks up significant amounts of moisture over the ocean. The PFE and HRSE humidity flux composite patterns show important distinctions.

Humidity flux at lower levels on the event day is shown in Fig. 8b. The humidity flux is averaged over 925-, 850-, and 775-hPa levels in the lower atmosphere. Figure 8b shows mean PFE humidity flux vectors (arrows) and the difference between the PFE and HRSE composites of the humidity flux component along the mean PFE wind direction (contours) on the event day. Arrows are shown only where the statistical significance of the PFE and HRSE difference is at least 99%.

The difference between the PFE and HRSE humidity flux composites is about 10% of the mean value. This difference persists over the 3 days leading up to the event.

Over continental India the humidity flux signal closely resembles the total mass flux patterns. The central and eastern parts are characterized by positive difference values, while negative values cover the western and southern parts. Unlike the pattern over the Arabian Sea, continental India values are the largest on the day of the event and considerably lower on previous days.

The persistent and higher-magnitude humidity flux reaching the western coast of India on the days before the PFEs, coupled with low surface pressure (Fig. 8a), gives rise to enhanced precipitation on the days leading to PFE events. The persistent precipitation prepares the surface wetness cofactor in producing the floods.

## 5. Conclusions

In this study, extended records of daily precipitation accumulation based on gauge stations are used in conjunction with unconventional flood reports to produce a concurrent proxy time series of potential flood events. Two precipitation datasets are used to assemble a century-long daily record over the Mumbai area. The period of overlap between the two precipitation records is analyzed to ensure consistency in extending the record. Since heavy precipitation and antecedent surface wetness are cofactors in producing floods, a simple time filter of the precipitation (the antecedent precipitation index) is used to provide estimates of surface conditions that, in combination with heavy precipitation, are likely to lead to flood events. The unconventional data—human reports of flood events—are used to define thresholds in the more objective instrumental records. The result is a century-long record of potential flood events over a region where stream discharge measurements to identify flood events are not easily available. For the period covering the twentieth century, nearly a thousand potential flood events and high rainfall and soil moisture events are identified. These are then used as the conditioning variable in broader area atmospheric field composites. The composites form statistical samples, used to identify the synoptic patterns that potentially lead to flood conditions over the study region. Back-trajectories are estimated for the composite samples to establish the sources and pathways of synoptic phenomena associated with both event types.

From 2 days before the event, PFEs are marked by persistent MSLP patterns whose minimum and, more

importantly, center location are distinct from non-flood but heavy precipitation events. The PFEs airmass flow pattern is confined by the Himalayan mountain range to the north, which enhances and intensifies the circulation. Over the Arabian Sea and peninsular India, PFE horizontal wind speeds at 850 hPa are about 2–4 $m\,s^{-1}$ stronger than in HRSEs samples. Higher differences are evident over northeastern continental India. PFE air particles that eventually reach the study area originate far afield over the waters between Madagascar and Somalia 2 days before the event. They follow a bowlike trajectory over the waters adjacent to the Horn of Africa and the Arabian Sea. Higher wind speeds at lower levels in the atmosphere enhance the moisture burden of the particle air masses. At 850 hPa, specific humidity increases as moist air passes over the Arabian Sea, with maximum differences over 65°E. At this latitude the PFE specific humidity ranges between 0.010 and 0.017, whereas HRSE composite samples have 850-hPa specific humidity values in the range from 0.008 to 0.014. These synoptic conditions correspond to active monsoon periods with embedded low pressure centers.

This study contributes to our understanding of atmospheric preconditions for flooding and hydrometeorological disaster events. The composites link the localized events to the larger synoptic conditions. The synoptic signature is useful for the short-term prediction of hydrometeorological hazards affecting the local region. Since the signatures are on the synoptic scale, they are better resolved than moist parameterized processes in atmospheric models used for prediction. Finally, the synoptic signatures can be used to assess the region's vulnerability to a changing climate. Again because large-scale synoptic atmospheric states are resolved and moist processes are parameterized, the changed frequency of the signature can be used to infer possible regional and local climate change impacts.

*Acknowledgments.* We are indebted to the European Centre for Medium-Range Weather Forecasts (ECMWF) for providing the ERA-40 reanalysis data. We thank MunichRe for providing the Indian Flood catalogues and discussions. We thank Sven Ulbrich (University of Cologne) for help with data handling and Andreas H. Fink (University of Cologne) for discussions. V. Venugopal (Indian Institute of Sciences, Bangalore) and Rajib Chattopadhyay (Indian Institute of Tropical Meteorology, Pune) kindly provided gridded precipitation data. Finally, we would like to thank the three anonymous reviewers for their helpful comments and suggestions, which contributed to improve the manuscript.

## REFERENCES

Annamalai, H., and J. M. Slingo, 2001: Active/break cycles: Diagnosis of the intraseasonal variability of the Asian summer monsoon. *Climate Dyn.,* **18,** 85–102, doi:10.1007/s003820100161.

Blanchard, B., M. McFarland, T. Schmugge, and E. Rhoades, 1981: Estimation of soil moisture with API algorithms and microwave emission. *J. Amer. Water Res. Assoc.,* **17,** 767–774, doi:10.1111/j.1752-1688.1981.tb01296.x.

Blanford, H. F., 1886: Rainfall of India. *Mem. Ind. Meteor. Dept.,* **2,** 217–448.

Bollasina, M., and S. Nigam, 2009: Indian Ocean SST, evaporation, and precipitation during the South Asian summer monsoon in IPCC-AR4 coupled simulations. *Climate Dyn.,* **33,** 1017–1032, doi:10.1007/s00382-008-0477-4.

Boos, W. R., and Z. Kuang, 2010: Dominant control of the South Asian monsoon by orographic insulation versus plateau heating. *Nature,* **463,** 218–222, doi:10.1038/nature08707.

Bowman, K., J. Collier, G. North, Q. Wu, E. Ha, and J. Hardin, 2005: Diurnal cycle of tropical precipitation in Tropical Rainfall Measuring Mission (TRMM) satellite and ocean buoy rain gauge data. *J. Geophys. Res.,* **110,** D21104, doi:10.1029/2005JD005763.

Bureau for Crisis Prevention and Recovery, 2004: Reducing disaster risk: A challenge for development. United Nations Development Programme, 10 pp. [Available online at http://www.un.org/special-rep/ohrlls/ldc/Global-Reports/UNDP%20Reducing%20Disaster%20Risk.pdf.]

Charney, J. G., 1969: The intertropical convergence zone and the Hadley circulation of the atmosphere. *Proc. Symp. on Numerical Weather Prediction,* Tokyo, Japan, WMO/IUCG, 73–79.

Gadgil, S., 2003: The Indian monsoon and its variability. *Annu. Rev. Earth Planet. Sci.,* **31,** 429–467, doi:10.1146/annurev.earth.31.100901.141251.

——, and P. V. Joseph, 2003: On breaks of the Indian monsoon. *Proc. Ind. Acad. Sci. (Earth Planet. Sci.),* **112,** 529–558.

Government of Maharashtra, 2005: Maharashtra floods 2005: Relief and rehabilitation. Division of Relief and Rehabilitation, Mumbai, Maharashtra, India, 17 pp. [Available online at http://mdmu.maharashtra.gov.in/pdf/Flood/statusreport.pdf.]

Hartmann, D. L., and M. L. Michelsen, 1989: Intraseasonal periodicities in Indian rainfall. *J. Atmos. Sci.,* **46,** 2838–2862, doi:10.1175/1520-0469(1989)046<2838:IPIIR>2.0.CO;2.

Hawksworth, J., T. Hoehn, and A. Tiwari, 2009: Which are the largest city economies in the world and how might this change by 2025? UK Economic Outlook, November, Price Waterhouse Coopers, 15–26.

He, J., C. Sun, Y. Liu, and J. Matsumoto, 2007: Seasonal transition features of large-scale moisture transport in the Asian-Australian monsoon region. *Adv. Atmos. Sci.,* **24,** 1–14, doi:10.1007/s00376-007-0001-5.

Izumo, T., C. de Boyer Montégut, J. Luo, S. K. Behera, S. Masson, and T. Yamagata, 2008: The role of the western Arabian Sea upwelling in Indian monsoon rainfall variability. *J. Climate,* **21,** 5603–5623, doi:10.1175/2008JCLI2158.1.

Kim, H., P. J. Webster, A. C. Curry, and V. E. Toma, 2012: Asian summer monsoon prediction in ECMWF system 4 and NCEP CFSv2 retrospective seasonal forecasts. *Climate Dyn.,* **39,** 2975–2991, doi:10.1007/s00382-012-1470-5.

Krishnamurthy, V., and R. S. Ajayamohan, 2010: Composite structure of monsoon low pressure systems and its relation to Indian rainfall. *J. Climate,* **23,** 4285–4305, doi:10.1175/2010JCLI2953.1.

Kummerow, C., W. Barnes, T. Kozu, J. Shiue, and J. Simpson, 1998: The Tropical Rainfall Measuring Mission (TRMM) sensor package. *J. Atmos. Oceanic Technol.,* **15,** 809–817, doi:10.1175/1520-0426(1998)015<0809:TTRMMT>2.0.CO;2.

——, and Coauthors, 2000: The status of the Tropical Rainfall Measuring Mission (TRMM) after two years in orbit. *J. Appl. Meteor.,* **39,** 1965–1982, doi:10.1175/1520-0450(2001)040<1965:TSOTTR>2.0.CO;2.

Ma, S., X. Rodó, Y. Song, and B. A. Cash, 2012: Dynamical linkage of tropical and subtropical weather systems to the intraseasonal oscillations of the Indian summer monsoon rainfall. Part I: Observations. *Climate Dyn.,* **39,** 557–574, doi:10.1007/s00382-012-1352-x.

Methven, J., 1997: Offline trajectories: Calculation and accuracy. Global Atmospheric Modelling Programme Tech. Rep. 44, University of Reading, 18 pp.

——, M. Evans, P. Simmonds, and G. Spain, 2001: Estimating relationships between air-mass origin and chemical composition. *J. Geophys. Res.,* **106,** 5005–5019, doi:10.1029/2000JD900694.

Munich Re Group, 2008a: Indian flood catalogue 1900–1947. Munich Re Group Internal Tech. Rep., Munich, Germany, 144 pp.

——, 2008b: Indian flood catalogue 1948–1978. Munich Re Group Internal Tech. Rep., Munich, Germany, 88 pp.

Oouchi, K., A. T. Noda, M. Satoh, B. Wang, S.-P. Xie, H. G. Takahashi, and T. Yasunari, 2009: Asian summer monsoon simulated by a global cloud-system-resolving model: Diurnal to intra-seasonal variability. *Geophys. Res. Lett.,* **36,** L11815, doi:10.1029/2009GL038271.

Ordóñez, P., P. Ribera, D. Gallego, and C. Peña-Ortiz, 2012: Major moisture sources for western and southern India and their role on synoptic-scale rainfall events. *Hydrol. Processes,* **26,** 3886–3895, doi:10.1002/hyp.8455.

Philander, S. G. H., D. Gu, G. Lambert, T. Li, D. Halpern, N. C. Lau, and R. C. Pacanowski, 1996: Why the ITCZ is mostly north of the equator. *J. Climate,* **9,** 2958–2972, doi:10.1175/1520-0442(1996)009<2958:WTIIMN>2.0.CO;2.

Rajeevan, M., J. Bhate, J. D. Kale, and B. Lal, 2006: High resolution daily gridded rainfall data for the Indian region: Analysis of break and active monsoon spells. *Curr. Sci.,* **91,** 296–306.

——, S. Gadgil, and J. Bhate, 2010: Active and break spells of the Indian summer monsoon. *J. Earth Syst. Sci.,* **119,** 229–247, doi:10.1007/s12040-010-0019-4.

Ranger, N., and Coauthors, 2011: An assessment of the potential impact of climate change on flood risk in Mumbai. *Climatic Change,* **104,** 139–167, doi:10.1007/s10584-010-9979-2.

Riehl, H., 1979: *Climate and Weather in the Tropics.* Academic Press, 611 pp.

Rodwell, M. J., 1997: Breaks in the Asian monsoon: The influence of Southern Hemisphere weather systems. *J. Atmos. Sci.,* **54,** 2597–2611, doi:10.1175/1520-0469(1997)054<2597:BITAMT>2.0.CO;2.

Sanderson, V., C. Kidd, and G. McGregor, 2006: A comparison of TRMM microwave techniques for detecting the diurnal rainfall cycle. *J. Hydrometeor.,* **7,** 687–704, doi:10.1175/JHM507.1.

Saxton, K., and A. Lenz, 1967: Antecedent retention indexes predict soil moisture. *J. Hydraul. Div.,* **93,** 223–241.

Uppala, S. M., and Coauthors, 2005: The ERA-40 Re-Analysis. *Quart. J. Roy. Meteor. Soc.,* **131,** 2961–3012, doi:10.1256/qj.04.176.

Vecchi, G. A., and D. E. Harrison, 2004: Interannual Indian rainfall variability and Indian Ocean sea surface temperature anomalies. *Earth's Climate: The Ocean–Atmosphere Interaction, Geophys. Monogr.,* Vol. 147, Amer. Geophys. Union, 247–259.

Wang, B., and H. Lin, 2002: Rainy season of the Asian-Pacific summer monsoon. *J. Climate,* **15,** 386–398, doi:10.1175/1520-0442(2002)015<0386:RSOTAP>2.0.CO;2.

Webster, P. J., V. O. Magana, T. N. Palmer, J. Shukla, R. A. Tomas, M. Yanai, and T. Yasunari, 1998: Monsoons: Processes, predictability, and the prospects for prediction. *J. Geophys. Res.,* **103,** 14 451–14 510, doi:10.1029/97JC02719.

Zhou, T. J., and R. C. Yu, 2005: Atmospheric water vapor transport associated with typical anomalous summer rainfall patterns in China. *J. Geophys. Res.,* **110,** D08104, doi:10.1029/2004JD005413.